Modeling Continuous Access EVA – Solution Performance

Don Fraser
Disaster Tolerant Solutions Architect
Continuous Access EVA Product Engineering
hp Network & Server Storage

HP WORLD 2003
Solutions and Technology Conference & Expo

Agenda

- Physics of distance
  - understanding why it takes so long
- Performance Estimator
  - creating an educated guess
  - based on
    - size of writes
    - distance between sites
    - link bandwidth
  - first a single write stream
  - then impact of multiple streams
- Questions taken throughout

Physics of Distance

Or why does it take so long?

- Speed of light is $3 * 10^8$ m per second in vacuum
  - in wire, 1 nano-second is 30 cm (12 inches)
- Speed of light is $2 * 10^8$ m per second in most fiber
  - in fiber, 1 nano-second is 20 cm (8 inches)
  - or 5 microseconds (μSec) per kilometer

Physics of Distance

Because, over distance it does!

- SCSI read/write and HSG80 replication is two round trips
  - each 100 km adds [5 μSec / Km *100 km * 4 trips] = 2 mSec
  - reduces performance similar to slower drive
    - 15k RPM -> 2 mSec average seek time
    - 7200 RPM -> 4 mSec average seek time
- Continuous Access EVA replication is one round trip
  - each 100 km adds 1 mSec
  - performance at 200 km is similar to using slower drives

Performance Estimator

Creating an educated guess

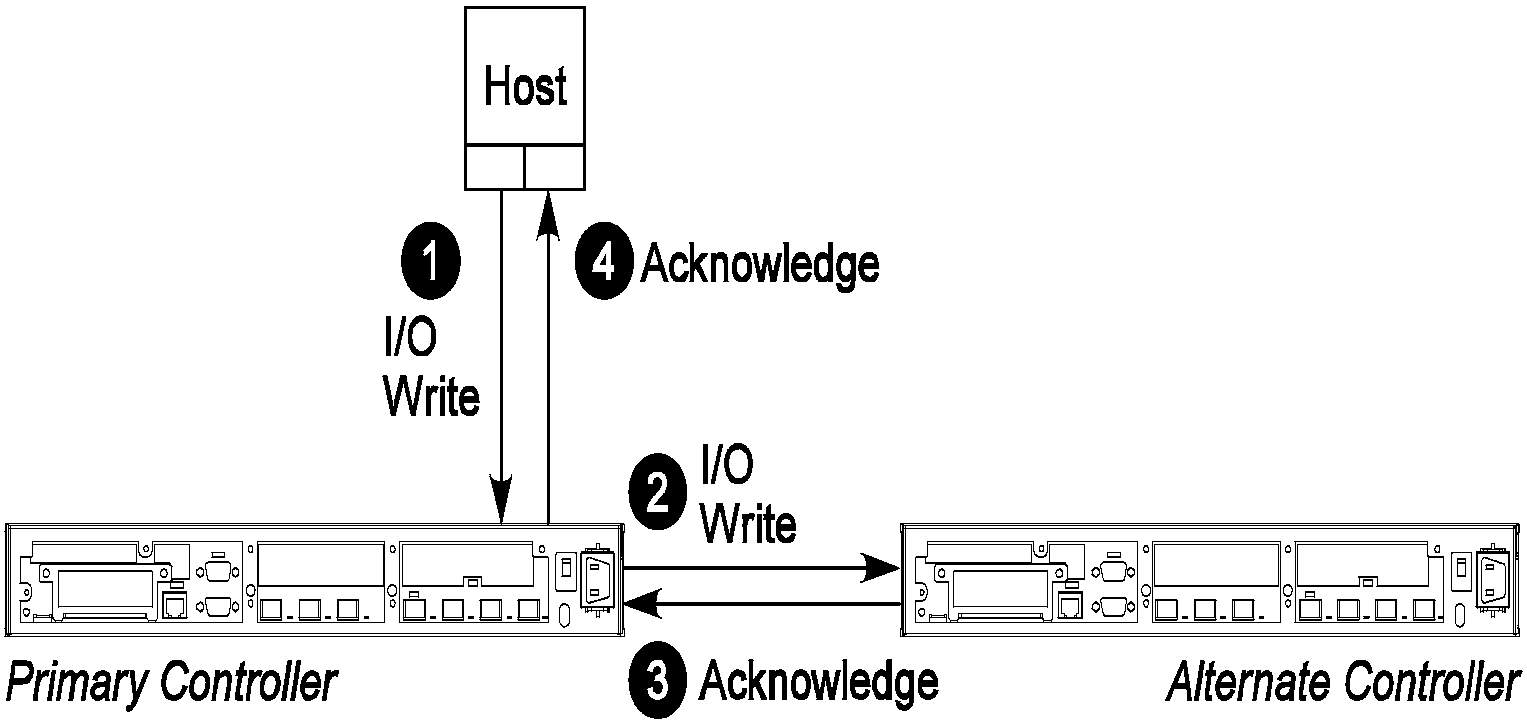
- Based on time to complete a single synchronous replication write for a given bandwidth of link
- Add impact of distance
- Calculate how many single writes per second
- Multiply by number of parallel streams
- Reduce by expected utilization

Performance Estimator

Time to complete a single synchronous replication at zero separation distance – 4 step process

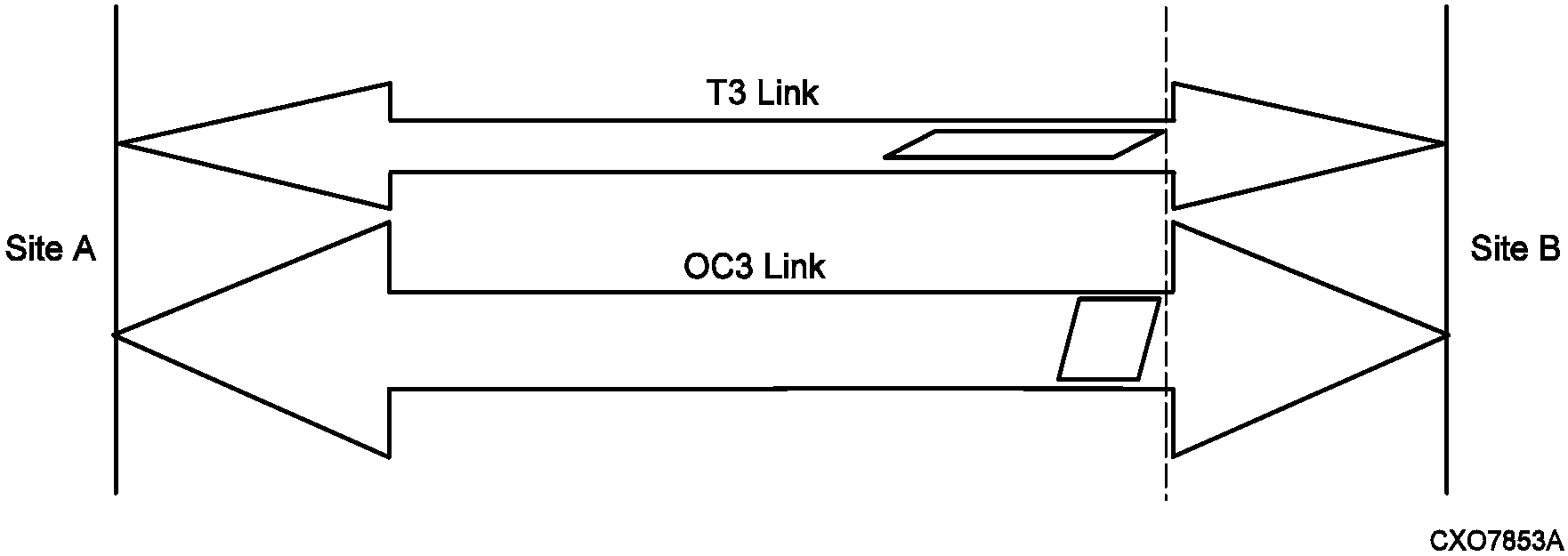
1. Host writes to first array's cache
2. Replicates the write, and sends it to second array
3. When in cache, second array acknowledges back to first array
4. First array acknowledges back to host issuing write.

Performance Estimator



Performance Estimator

- **Effects of distance does not depend on intersite link bandwidth**



- **Delay is 5 microsecond per kilometer per trip**

Performance Estimator

Calculate number of single writes per second

- Invert response time in mSec per Write
- Yields peak writes per second
- For a single I/O stream
  - a single, simple application

Using the Performance Estimator

hp StorageWorks Continuous Access EVA Replication Performance Estimator - V1, VCS V3
For a Single I/O Stream using direct connect fiber, WDM, or Fiber Channel over IP
Estimates based on One Outstanding Synchronous Replication of Local Writes

Enter One Way Intersite Latency: **2.0** ms — **400** km or **249** miles

Enter Size of a Write Data Packet (KiloBytes): **8** KB — 256 KB max.

	2 Gbps fiber	1 Gbps fiber	FCIP	FCIP
Link Bandwidth in megabits / sec	2000 Mbps	1000 Mbps	100.0 Mbps	10.0 Mbps
Packet Load/ Unload Time:	0.21 ms	0.27 ms	1.01 ms	7.18 ms
ms per I/O:	4.56 ms	4.66 ms	5.38 ms	13.85 ms
I/Os per Second:	219.5 I/O/sec	214.4 I/O/sec	185.9 I/O/sec	72.2 I/O/sec
or	6.32 GB/h	6.17 GB/h	5.35 GB/h	2.08 GB/h
Approximately:	17.56 Mbps	17.15 Mbps	14.87 Mbps	5.78 Mbps
% Bandwidth	0.88%	1.72%	14.87%	57.77%

Performance Estimator

- Estimator only tells you how many writes per second are available using a single I/O stream – a first guess

- Now look at how to increase that for multi-stream I/O or multi-threaded applications

Performance Estimator

- Start with knowing controller limits
  - each replication path has 32, 8KB buffers
  - writes up to 8K use one buffer
  - writes over 8K use more than one buffer
  - port 1 to port 1 is one path, port 2 to port 2 another
- Also understand that link utilization will limit bandwidth
  - don't over subscribe
  - limits are: average 40%; peak 45%
- Finally the controller can only do so much over time.
  - current understanding is x6 at zero distance
  - increases as distance allows for more outstanding writes

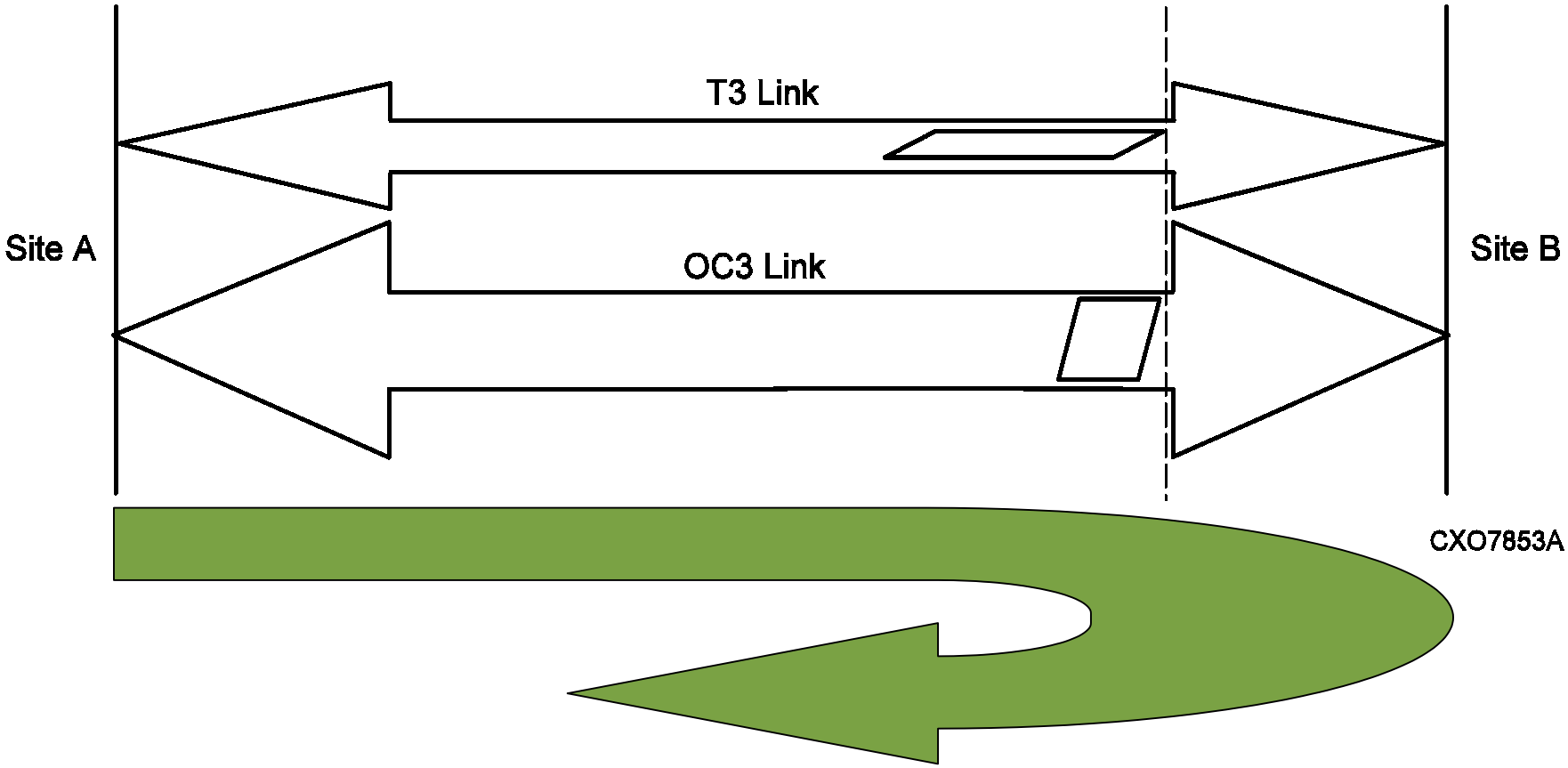
Performance Estimator

Reduce by expected utilization

- Use 100 % only if dedicated environment and able to issue writes as soon as previous completes
- Use 70% as theoretical peak, 50% as practical peak
- Remember Ethernet
  - a full one only uses 50% of capacity on average
  - and 70% utilization is not seen

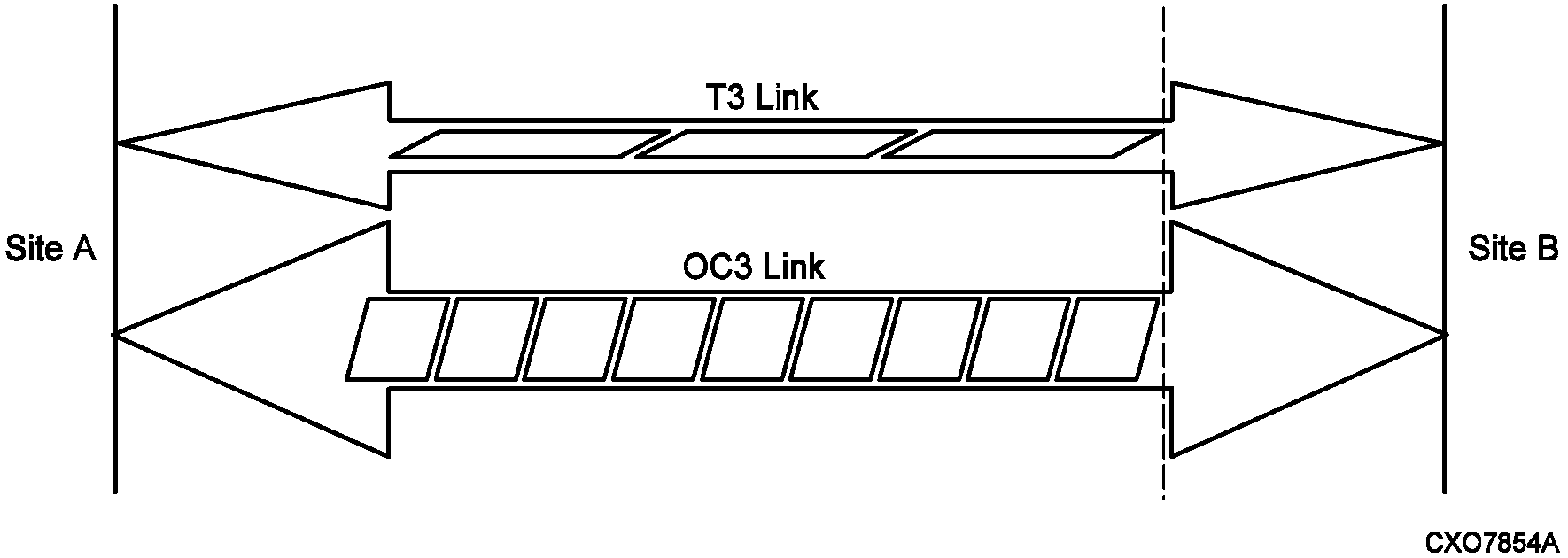
Performance Estimator

- *First estimate is based on number of single synchronous writes per second*



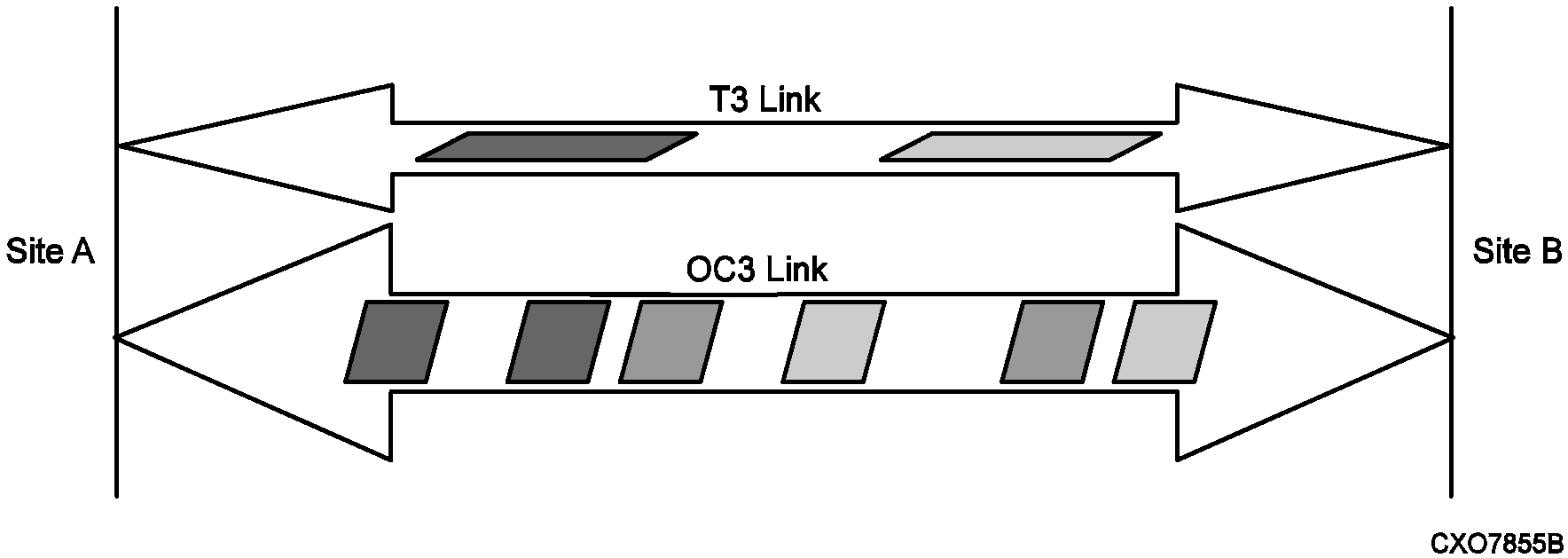
Performance Estimator

Then estimate effect of streaming multiple I/Os for a single application at 100 % expected utilization



Performance Estimator

Finally consider effect of streaming multiple I/Os for a single application at 70 % expected utilization



Performance Estimator

In review, it's about creating a good guess!

- Know the size of the write and the distance (delay)
- Figure out how fast that one could be repeated
- Understand impact of multiple I/O in the pipe
  - 70 % peak
  - 50% expected
  - 40% average, 45% peak for planning purposes
- Understand impact of limits
  - link bandwidth
  - controller bandwidth

For more information

- Documents mentioned in this presentation are available from the Continuous Access EVA web site

  http://h18006.www1.hp.com/products/storage/software/conaccesseva/index.html

  - And then click on "technical documentation". As a starting point, see the Continuous Access EVA Design Guide.

Interex, Encompass and HP bring you a powerful new HP World.